ISO 191** Standards for NASA

SiriJodha Singh Khalsa, Lead

MENDS Tiger Team

Motivation

- Movement within NASA to support ISO metadata
  - GCMD and ECHO both actively working on ISO
  - Missions and data centers looking at ISO
- FGDC has endorsed ISO 19115
  - OMB mandated all Federal agencies to supply metadata for their geospatial datasets in FGDC-approved standard
- DAAC Managers Meeting in May recommended
  - A team of experts and stakeholders be formed to study and make recommendations regarding ISO 19115

Team Composition & Collaboration

- Supported by ESDIS Project
- Representatives from
  - Clearinghouses: ESDIS, ECHO, GCMD, and ECS
  - Data Centers: GES DISC, NSIDC, GHRC, PO.DAAC, ORNL, ASF; also NOAA NGDC
  - Missions: SMAP, GHRSST, OIB
- Collaboration via team website, email list and web meetings
- Acronym MENDS: Metadata Evolution for NASA Data Systems

Goals

- Determine how NASA Earth Science Data Systems should support ISO 19115 metadata
  - Weighing costs and benefits
  - Consider mission data systems, data archive and distribution systems, and metadata catalogs and clearinghouses (ECHO and GCMD)
- Coordinate ISO efforts within NASA ES data systems
- Determine desired level, and meaning of compliance
- Study ways to take full advantage of ISO
  - Investigate advantages, benefits, pitfalls
- Provide input to revision, 19115:2012
  - Determine best avenue for providing NASA participation
  - Define clear process for NASA's single voice

Approach

- Analyze stakeholder needs and constraints
- Assemble library of reference material
- Study current implementations of ISO within NASA
- Evaluate existing profiles (NAP, ANZ, WMO, INSPIRE,…)
- Seek advice from other agencies (i.e. NOAA)
- Develop recommendations regarding:
  - Whether ISO should be a requirement for DS missions and if so, what this means
  - Whether NASA should simply adopt the NAP, develop its own ISO profile or just define best practices
  - Optimal solutions for ECHO, missions, data centers

Schedule

- Phase I (Aug – Dec, 2010)
  - Gather information, evaluate, make recommendations
  - Draft document expected early December
  - Presentation at Fall AGU
  - Wrap-up, Phase II kickoff at Winter ESIP Federation Meeting (Jan 2011)
- Phase II (6 mo.)
  - Develop schema/profile/best practices of 19115
- Phase III (12 mo.)
  - Develop guidelines, tools, and best practices

Background on 19115

- Formal title is 19115:2003 Geographic Information – Metadata
- It is an *abstract* model for metadata
  - Another standard, ISO 19139:2007 specifies the XML implementation of 19115
- ISO 19115 defines metadata elements
  - Provides a schema (in UML)
  - Gives definitions and constraints (in a data dictionary)

ISO 19115 Does Not Stand Alone

- UML Packages used in 19115 (w/ corrigendum)
  - CI Citation (ISO 19115)
  - DQ Data quality (ISO 19115)
  - DS Dataset (ISO 19115)
  - EX Extent (ISO 19115)
  - GF General Feature (ISO 19109)
  - GM Geometry (ISO 19107, replaced by GML in ISO/TS 19139)
  - LI Lineage (ISO 19115)
  - MD Metadata (ISO 19115)
  - RS Reference System (ISO 19115)
  - SC Spatial Coordinates (ISO 19111)
  - SV Services (ISO 19119)
  - TM Temporal Schema (ISO 19108)

Plus...

- 19115-2 additions
  - LE Lineage Extended (ISO 19115-2)
  - MI Metadata for Imagery (ISO 19115-2)
  - MX Metadata – XML schema (ISO/TS 19139)
  - QE Data Quality Extended (ISO 19115-2)

In Addition…

- Also relevant:
  - CV Coverages (ISO 19123)
  - FC Feature Catalogue (ISO 19110)
  - FE Feature (ISO 19109)
  - PF Feature Portrayal (ISO 19117)
  - **SD Sensor Data** (ISO/TS 19130)

Extensions and Profiles

- ISO 19115 can be extended by adding new:
  - Sections (packages)
  - Codelists
  - Codelist elements
  - Metadata entities (classes)
  - Metadata elements (class attributes)
- A Community Profile of ISO 19115 can include
  - Extensions
  - Additional constraints
  - Must include core

Stakeholder Survey

- Current usage of metadata
- Current ISO capabilities, if any
- Available ISO expertise and tools
- Constraints and directives
- Impediments and concerns
- Future plans and available resources

Categories of Metadata Usage

- Spatial and Temporal extent
- Spatial and Temporal representation (CRS, grid resolution, date/time encoding, etc.)
- Platform and sensor specifications
- References to documentation and responsible parties
- Product lineage
- Production version information, algorithmic parameters
- Data quality information
- Services available on the data
- Access constraints
- Data granule characteristics
- Additional descriptive information (collection and granule)

Mission uses of Metadata

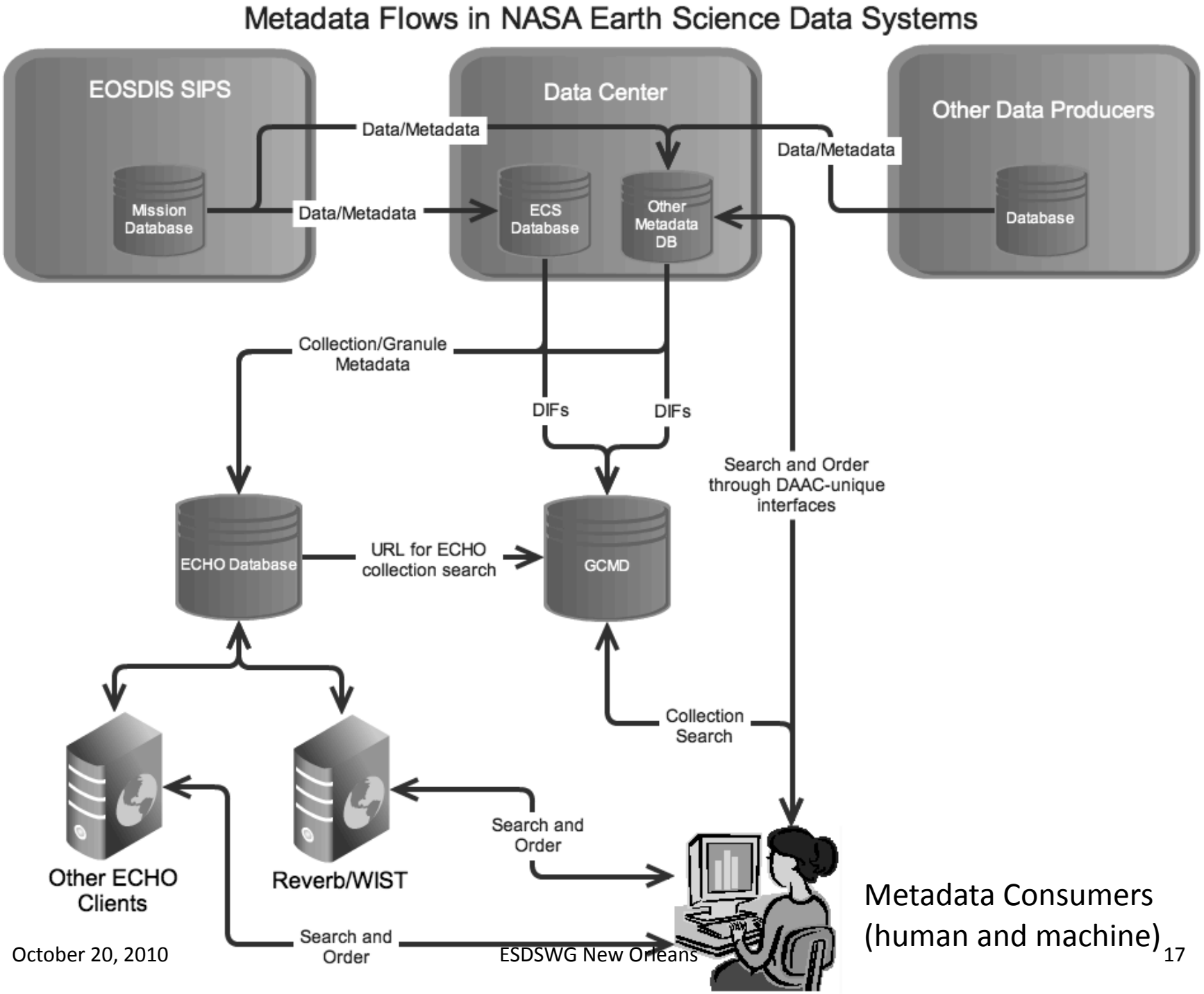
- To properly describe acquired data to facilitate search and discovery
- To support data usage inside an outside of NASA
- Exporting to data centers for archiving and other data management tasks

Data Center uses of Metadata

- Data search and discovery
- Data subsetting and other services
- Data management, order processing and tracking
- Constructing catalog pages
- Exporting to clearinghouses and to consumers in multiple formats

Clearinghouse uses of Metadata

- Support data search and discovery
- Construct catalog pages
- Export metadata to consumers in multiple formats


Metadata Flows in NASA Earth Science Data Systems
EOSDIS SIPS
Mission Database
Data/Metadata
Data/Metadata
Data Center
ECS Database
Other Metadata DB
Data/Metadata
Other Data Producers
Database
Collection/Granule Metadata
DIFs
DIFs
Search and Order through DAAC-unique interfaces
ECHO Database
URL for ECHO collection search
GCMD
Collection Search
Other ECHO Clients
Reverb/WIST
Search and Order
Search and Order
Metadata Consumers (human and machine)

Some Issues

- Complexity of the ISO standard(s)
  - Steep learning curve
  - Current lack of expertise
- Availability of the ISO standard(s)
  - Copyrighted, not easy to access
  - Availability of tools could mitigate
- Minimizing impact on existing systems
  - Cost: Staff, tools, reconfiguration

More Issues

- Common usage across NASA ESDS highly desirable
  - For example, where to place ShortName?
- What should be searchable?
  - Expand what ECHO supports?
- Being responsive to user community needs
  - Can NASA influence evolution of the standard?

Important Considerations

- Should ISO be a requirement for DS missions?
  - Or are ECHO/DIF mappings sufficient?
- If NASA decides to recommend a profile of ISO should it:
  - Adopt the NAP
  - Develop its own profile
  - Define best practices
- Possible paths for adoption of ISO by Data Centers, ECS, ECHO, future missions
  - Respect timelines
  - Account for comparative impacts on stakeholders

To be Continued…
in Tomorrow's SPG Breakout

Details…

QUESTIONS?

BACKUP SLIDES

ISO 19115 Recommended Core

Metadata elements answering the following questions: “Does a dataset on a specific topic exist (‘what’)?”, “For a specific place (‘where’)?”, “For a specific date or period (‘when’)?” and “A point of contact to learn more about or order the dataset (‘who’)?”.

- ➳ ***Dataset title***
- ➳ ***Dataset reference date***
- — Dataset responsible party
- ➳ ***Geographic location of the dataset*** **(by four corners or by geographic identifiers)**
- ➳ **Dataset language**
- ➳ **Dataset character set**
- ➳ ***Dataset topic category***
- – Spatial Resolution
- ➳ ***Abstract describing the dataset***
- – Distribution format

- – Additional extent information (vertical and temporal)
- – Spatial representation type
- – Reference system
- – Lineage statement
- – On-line resource
- – Metadata file identifier
- – Metadata standard name
- – Metadata standard version
- ➳ ***Metadata language***
- ➳ ***Metadata character set***
- ➳ ***Metadata point of contact***
- ➳ ***Metadata time stamp***

➳ ***Mandatory***

System Supplied